THEMIS SIGNAL ANALYSIS STATISTICS RESEARCH PROGRAM

APPROXIMATIONS TO LARGE PROBABILITIES OF ALL SUCCESSES FOR

GENERAL CASE AND SOME OPERATIONS RESEARCH IMPLICATIONS

by

John E. Walsh and C. J. Kelleher

Technical Report No. 25
Department of Statistics THEMIS Contract

February 20, 1969

DEPARTMENT OF STATISTICS
Southern Methodist University

$$p \ \cdots \ (\cdots (i-1))$$

APPROXIMATIONS TO LARGE PROBABILITIES OF ALL SUCCESSES FOR

GENERAL CASE AND SOME OPERATIONS RESEARCH IMPLICATIONS

John E. Walsh
Southern Methodist University[*]
Dallas, Texas, U.S.A.

G. J. Kelleher
Institute for Defense Analyses
Arlington, Virginia, U.S.A.

ABSTRACT

There are many cases where an overall effort is successful if and only if the efforts (or events) of a sequence are all successful. Often, the principal interest is in cases where overall success has a large probability (say, at least .8). Suppose there are n efforts in the sequence and that $p_i(s_{i-1})$ is the probability that the i-th effort is a success given that preceding efforts $1, \cdots, i-1$ are successes $(i=1, \cdots, n)$, where s_o denotes no conditions. An approximate value, also sharp upper and lower bounds, are developed for the probability that all n events are successes. This is done for various levels of generality, including a form of complete generality. These results depend only on n, the generality level, and the arithmetic average of the $p_i(s_{i-1})$. They are useful when the probability of all successes is at least .8; then the approximate value is near both bounds. The necessity of only considering the arithmetic average of the $p_i(s_{i-1})$, rather than their product, sometimes can be useful in analyses of an operations research nature (including reliability situations). Consider optimum allotment of resources to obtain a stated high probability of all successes. This can be obtained by minimizing the resource use subject to the arithmetic average of the $p_i(s_{i-1})$ equaling a determined value. This minimization is often less complicated than minimization subject to the product of the $p_i(s_{i-1})$ having the stated value. Also, statistical estimation of the probability of all successes is simplified when the expression using the arithmetic average of the $p_i(s_{i-1})$ is considered.

INTRODUCTION AND RESULTS

Consider n binomial events (not necessarily independent) where success and failure are the possible outcomes for an event. Often, there is interest in whether all the events are successes. That is, there is overall success if and only if every event results in success. In fact, many kinds of overall efforts can be considered to occur as a sequence of steps (or events) with overall success occurring if and only if all of the steps result in success. This is often the case for reliability situations and for accomplishment of missions (for example, military missions).

The principal interest is frequently in high probabilities of overall success (say, at least .8). Only cases with high probabilities are considered here.

Let the n events be numbered according to the sequence in which they can be considered to occur (event 1 is resolved first, etc. Use $p_i(s_{i-1})$ to denote the probability that the i-th event results in success given that events 1, $\cdots$, i-1 are successes (i = 2, $\cdots$, n), while $p_1(s_0) = p_1$ is the unconditional probability that the first event is a success. Then,

$$P(\text{all successes}) = \prod_{i=1}^{n} p_i(s_{i-1})$$

is the probability that all the events are successes

If the approximation is sufficiently accurate, there are advantages in

expressing P(all successes) in terms of the arithmetic average (denoted by p)

of the $p_i(s_{i-1})$. First, consider operations research type situations where

a stated large value of P(all successes) is required and this is to be obtained

with an optimum allotment of resources. This can be accomplished (approximately)

by minimizing the resource use subject to setting p equal to a determined

value. Experience indicates that this minimization is usually less complicated

than minimization subject to the product of the $p_i(s_{i-1})$ having the stated

value.

A second advantage is in performing statistical investigation when the

$p_i(s_{i-1})$ must be estimated. Probabilistic properties of an arithmetic average

of estimates are usually much more easily determined than those of a product

of estimates.

Estimation of a $p_i(s_{i-1})$ is often not difficult. That is, $p_i(s_{i-1})$

is ordinarily the probability that the i-th event is a success given that

nothing very unusual has happened (something very unusual happens only if at

least one of events 1, $\cdots$, i-1 is a failure). For example, a failure in a

reliability situation can result in a large stress (perhaps an explosion)

in the overall system. No unusual stress occurs if all preceding events are

successes. Thus, effectively, $p_i(s_{i-1})$ can be estimated by observations

on the i-th event when it receives no unusual stress. Fortunately, this is

the case in which observations are most easily and inexpensively obtained.

In fact, it is often possible to separately consider the i-th event (per-haps removed from the overall system) if suitable environmental conditions (temperature, pressure, etc.) are maintained.

For the cases considered, the level of generality is defined in terms of the number r of the $p_i(s_{i-1})$ are required to be at most equal to p. Complete generality occurs for $r = 1$ and the generality level decreases as r increases. However, r values as large as, say, $n/4$ would still seem to represent a moderate degree of generality.

Only cases where n, p, r are such that $(n-r+1)(n-p)$ is less than unity are considered. Then, for any meaningful n and r,

$$[1 - (n-r+1)(1-p)]p^{r-1} \le P(\text{all successes}) \le p^n .$$

The lower bound is at least .8 if $n(1-p) \le .2$ and is a monotonically increasing function of r. The bounds are close together and P(all successes) is very nearly equal to

$$(1/2)p^{r-1}[p^{n-r+1} + 1 - (n-r+1)(1-p)]$$

when $[1 - (n-r+1)(1-p)]p^{r-1}$ is at least .8.

The final section contains derivations of these probability results. When the $p_i(s_{i-1})$ are all very large, n is large, and dependence is of a restricted nature, the more easily applied approximate results of

(Walsh, 1955) are usable.

VERIFICATION

For notational simplicity, let $p_i = p_i(s_{i-1})$. Since the geometric mean of the p_i is at most equal to their arithmetic mean, P(all successes) $\leq p^n$, with equality possible (when the p_i are equal). This upper bound and its basis are applicable for all r.

Now, consider derivation of the sharp lower bound. The value of P(all successes) is

$$\prod_{i=1}^{n} p_i = \exp \left\{ \sum_{i=1}^{n} \log_e [1 - (1 - p_i)] \right\}$$

$$= \exp \left[-\sum_{i=1}^{n} \sum_{j=1}^{\infty} (1 - p_i)^j / j \right]$$

$$= \exp \left[-\sum_{j=1}^{\infty} j^{-1} \sum_{k=0}^{j} \binom{j}{k} (1 - p)^{j-k} \sum_{i=1}^{n} (p - p_i)^k \right]$$

For $(n-r+1)(1 - p) < 1$ and $k \geq 2$, $\sum_{i=1}^{n} (p - p_i)^k$ is largest when all but r of the p_i are unity, r-1 of them equal p, and the other one is such that their arithmetic average is p (easily verified by considering the cases of $k = 2, 3$ and the relationships for larger k). This implies that this other p_i is $1 - (n-r+1)(1 - p)$, so that the maximum is

$$\sum_{i=1}^{n} (p - p_i)^k = (n - r)^k (1 - p)^k + (-1)^k (n - r)(1 - p)^k \ .$$

Thus, since $r-1$ of the p_i are equal to p and also

$$\sum_{k=0}^{j} \binom{j}{k} [(n - r)^k + (-1)^k (n - r)] \quad [(n - r) + 1]^j + (n - r)(1 - 1)^j \ ,$$

which equals $(n-r+1)^j$, P(all successes) is at least equal to p^{r-1} times

$$\exp \left\{ - \sum_{j=1}^{\infty} j^{-1} (1 - p)^j \sum_{k=0}^{j} \binom{j}{k} [(n - r)^k + (-1)^k (n - r)] \right\}$$

$$= \exp \left\{ - \sum_{j=1}^{\infty} j^{-1} [(n-r+1)(1 - p)]^j \right\}$$

$$= \exp \left\{ \log_e [1 - (n-r+1)(1 - p)] \right\} - 1 - (n-r+1)(1 - p) \ ,$$

with equality possible.

REFERENCE

Walsh, John E., "The Poisson distribution as a limit for dependent binomial

events with unequal probabilities," Operations Research, Vol. 3 (1955),

pp. 198-209.

DOCUMENT CONTROL DATA · R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1 ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
SOUTHERN METHODIST UNIVERSITY	UNCLASSIFIED
	2b. GROUP
	UNCLASSIFIED

3 REPORT TITLE

Approximations to Large Probabilities of All Successes for
General Case and Some Operations Research Implications

4 DESCRIPTIVE NOTES (Type of report and inclusive dates)
Technical Report

5 AUTHOR(S) (First name, middle initial, last name)

John F. Walsh and G. J. Kelleher

6 REPORT DATE	7a. TOTAL NO OF PAGES	7b. NO OF REFS
February 20, 1969	6	1
8a CONTRACT OR GRANT NO	9a. ORIGINATOR'S REPORT NUMBER(S)	
N00014-68-A-0515		
b PROJECT NO	25	
NP 042-260		
	9b OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
d.		

10 DISTRIBUTION STATEMENT

Distribution of this document is unlimited.

11 SUPPLEMENTARY NOTES	12 SPONSORING MILITARY ACTIVITY
	Office of Naval Research

13 ABSTRACT

There are many case where an overall effort is successful if and only if
the efforts (or events) of a sequence are all successful. Often, the
principal interest is in cases where overall success has a large probability
(say, at least .8). Suppose there are n efforts in the sequence and that
$p_i(s_{i-1})$ is the probability that the i-th effort is a success given that
preceding efforts $1, \ldots, i-1$ are successes $(i=1, \ldots, n)$, where s_0 denotes
no conditions. An approximate value, also sharp upper and lower bounds, are
developed for the probability that all n events are successes. This is
done for various levels of generality, including a form of complete gener-
ality. These results depend only on n, the generality level, and the
arithmetic average of the $p_i(s_{i-1})$. They are useful when the probability of

all successes is at least .8; then the approximate value is near both bounds
The necessity of only considering the arithmetic average of the $p_i(s_{i-1})$,

rather than their product, sometimes can be useful in analyses of an operations
research nature (including reliability situations). Consider optimum allot-
ment of resources to obtain a stated high probability of all successes. This
can be obtained by minimizing the resource use subject to the arithmetic average
of the $p_i(s_{i-1})$ equaling a determined value. This minimization is often less
complicated than minimization subject to the product of the $p_i(s_{i-1})$ having the
stated value. Also, the estimation of all successes is simplified.

DD FORM 1473